

38188-382

*Patent*

UNITED STATES PATENT APPLICATION

FOR

SYSTEM AND METHOD FOR COMPOSITE CUSTOMER SEGMENTATION

INVENTOR:

ALAN K. GORENSTEIN

PREPARED BY:  
McDERMOTT, WILL & EMERY  
600 13<sup>TH</sup> STREET, N.W.  
WASHINGTON, DC 20005  
(202) 756-8600

*CUSTOMER*  
SYSTEM AND METHOD FOR COMPOSITE CONSUMER SEGMENTATION

## FIELD OF THE INVENTION

The present invention relates to customer marketing methods and more particularly, to strategies for segmenting customers and potential customers to increase efficiency of marketing efforts.

## BACKGROUND OF THE INVENTION

Marketing programs can include, for example, mail and direct mail campaigns, inbound and outbound telemarketing campaigns, and inbound and outbound web and e-mail campaigns.

In the field of customer targeting, a single segmentation method is used to attempt to select those customers who are most likely to respond to marketing programs. This initial segmentation is sometimes, but rarely, followed by a second independent segmentation to determine which customers are most likely to spend more if they respond. This second segmentation of customers is used primarily to reduce quantities of targeted customers contacted in order to meet a pre-specified criteria.

Segmentation strategies well known to one in the art can include, for example, linear models, logistic models, RFM segmentation, and CHAID (Chi-square Automatic Interaction Detection) segmentation.

Conventional wisdom in the marketing industry recognizes that each of the segmentation strategies perform better than the other strategies in certain situations depending on a number of various circumstances. Accordingly, through past experience, or by running different trials, the segmentation strategy is typically selected which optimizes, or emphasizes, the differentiation between the customer population for a desired outcome. The other, unselected segmentation strategies are then ignored because they are considered to be weaker indicators of the variance within the customer population. Within the framework of the selected segmentation strategy,

additional improvements to the results can be attempted but usually only by adding additional variables to be considered during the model analysis.

In the 1930s, RFM segmentation was developed. This method of segmentation sorts customers by the Recency of their last purchase, then by the Frequency of their purchases recorded on file, and finally by the Monetary value of their purchases recorded on file. RFM segmentation, and variations thereof, are still the primary methods used today by marketers to segment customers.

With the advent of computer technology and automation, companies have begun to shift to regression based segmentation methods. These relatively new methods for segmentation involve creating variables based on customers' purchases and demographic data. Next a specific event is targeted, or identified, (e.g., the likelihood of purchasing from a particular catalog) and then, the regression is run to create a statistical model that attempts to predict the targeted event. Almost always, regression methods of segmentation provide better results than RFM segmentation.

But, if RFM segmentation produces, for example, response rates of 3%, and running a regression produces response rates of 4%, then there is still room for improving response rates to the remaining 96% of customers and potential customers that have had marketing material remitted to them. These contacted but non-responding groups represent the bulk of the expenses involved in marketing today.

## SUMMARY OF THE INVENTION

While there is some value to the individual results of each conventional segmentation method, by combining multiple segmentation strategies, a synergistic effect can be realized. The value of the combined strategies is greater than any one of the independent views, resulting in consistently higher returns on marketing investments.

The present invention allows for the combination of any and all existing, and future, segmentations that independently are designed to explain variance. The inventive process yields higher marketing response rates and revenues per sale, while simultaneously allowing for lower marketing costs by reducing submission of marketing materials to unprofitable segments.

While the present invention is introduced and explained within the environment of marketing, this environment is merely exemplary and the broader concepts of the invention have applicability in any field, such as insurance and credit risk analysis, where improved customer segmentation can yield improved results and efficiencies.

Additional needs, advantages, and novel features of the present invention will be set forth in the description that follows, and in part, will become apparent upon examination or may be learned by practice of the invention. The features and advantages of the invention may be realized and obtained by means of the instrumentalities and combinations particularly pointed out in the appended claims.

#### BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is illustrated by way of example, and not by way of limitation, in the figures of the accompanying drawings and in which like reference numerals refer to similar elements and in which:

FIG. 1 illustrates a flowchart of a process for segmenting customers according to embodiments of the present invention.

FIG. 2 illustrates an exemplary computer platform on which an embodiment of the present invention may be implemented.

#### DESCRIPTION OF THE PREFERRED EMBODIMENT

The process flow illustrated in FIG. 1 depicts an exemplary method for improving customer segmentation for marketing purposes according to an embodiment of the present invention.

The exemplary flowchart of FIG. 1 begins, in step **S102**, with the creation of different, independent segmentation strategies. These segmentation strategies can include conventional strategies such as linear models, logistic models, RFM segmentation, CHAID segmentation, CART segmentation, etc. The different strategies can be used to independently predict a single targeted event.

Also, each single strategy can be used to predict distinct targeted events, even though the names that are selected will be used for a single pre-determined goal. For example, if two logistic processes are selected, one logistic model might target "the likelihood of responding to marketing program X" and another logistic model might target "the likelihood of responding to any marketing program during the coming year".

After the multiple segmentation strategies are created, each of the independent segmentation strategies is evaluated, in step S104, in a corresponding lift table to determine which segments of the table independently provide an acceptable level of predicted profitability.

A lift table is a chart that attempts to explain how a proposed segmentation strategy will function when it is actually used in the marketplace. In direct marketing, a lift table would show the estimated response rate, estimated average revenue per marketing piece mailed, and the estimated average value of each transaction for each segment. Each segment would be created by viewing individual ranks, or by creating groups of ranks (e.g., based on percentiles, deciles, etc.). Using a lift table, those segments which are more likely to be profitable, and by what potential degree they are likely to be profitable, can be determined.

Some or all of the independent segmentation strategies are then performed, in step S106, against a list of customers or potential customers to determine a score for each customer for each of the different targets (i.e., segmentation strategies). Unlike conventional marketing practices, results from non-optimal segmentation strategies are not discarded but, instead, are retained and used as meaningful and significant data.

The score that each customer receives from each independent segmentation strategy is determined and then stored. Many functionally equivalent methods of storing these scores, and any of the data structures described herein, can be used. Relational as well as object oriented database management systems, for example, can be used that execute on a single platform, or on a distributed computing platform. Similarly, the scores, and other data, can be stored local to the computing platform performing the segmentation strategies and analysis, or can be stored remote therefrom.

The particular format, and ranges, of the scores depend on the segmentation strategy and are considered as representative of the target variable of the strategy. For example, if the target of a segmentation strategy is "the predicted spending from catalog X", then \$15.00 may be the mean predicted score. Accordingly, the scores for that strategy will be in units of dollars and their significance can be measured by their difference from \$15.00. If the target variable involves "a likelihood of Y occurring", then the scores may be viewed as percentages and ranked, or sorted, accordingly.

Within each segmentation strategy, each customer is ranked, in step S108, based upon each score that has been received and recorded from that particular independent segmentation strategy. The rank assigned to a customer is, as conventionally known, a natural number indicating that customer's location in an ordered list.

Even though some of the segmentation strategies may, individually, be less optimal than others (and according to conventional wisdom, therefore, should not be considered), the present invention considers the results of different, independent segmentation strategies by combining them to generate a composite score. Such a combination of segmentation strategies can be performed in one of the following, exemplary ways or by combining two or more of the exemplary ways simultaneously or sequentially.

One way to combine the different segmentation strategies is to generate, in step S110, an average score. Each customer's ranks from the different strategies are combined in an average which results in a "rank-based score" which is, itself, then ranked. The average that is calculated can be a straight average or a weighted average based upon, for example, a relative value, or scaling factor, associated with each segmentation strategy.

Another way to combine the different segmentation strategies is to determine, in step S112, a customer's value based on some evaluation of that customer's ranking within the different segmentation strategies. For example, if a customer is in the top rank for three independent segmentation strategies and is in the bottom rank for a fourth strategy, that customer might still be considered to have a relatively high value, and therefore, selected (even if the customer's overall

average rank, determined above, would have been considered "poor"). Conversely, if a customer is in the top rank for one independent segmentation strategy, the second rank for two independent segmentations, and in a bottom rank for the fourth, that customer may be considered to have a relatively low value, and therefore, not selected (even though the customer's overall average rank, determined above, would have been considered "average").

The levels, or customer values, of what are acceptable rank combinations, and what are not, can be extrapolated from the lift tables established in prior steps, or through an ANOVA (Analysis of Variance Between Groups). Using this analysis, all of the rank combinations for a customer are assigned values and, from these values, a new consolidated ranking for each customer is assigned by determining which combinations are most likely to predict events.

The test used in an ANOVA compares the variation (measured by the variance) *between* populations with the variation *within* populations. If the "between variation" is much larger than the "within variation", the means of the different populations will not be equal. If the between and within variations are approximately the same size, then there will be no significant difference between population means.

This procedure employs the F-statistic, or F-value, to test the statistical significance of the differences among the obtained means of two or more random samples from a given population. More specifically, using the Central Limit Theorem, one calculates two estimates of a population variance.

The F-statistic is formed as the ratio of these two estimates. If this ratio is sufficiently larger than 1, the observed differences among the obtained means are described as being statistically significant.

Within the exemplary marketing environment herein described, the F-statistic can be considered as a composite, or consolidated, score that takes into account all the different segmentation strategies. A ranking of the customers based on this composite score can then be performed.

For example, if there are four segmentation strategies, and each segmentation strategy has 100 ranks, five exemplary combinations of ranks could be:

Customer Identifier	Segmentation Strategy #1	Segmentation Strategy #2	Segmentation Strategy #3	Segmentation Strategy #4
1	01	23	34	32
2	94	12	12	43
3	15	52	19	62
4	83	18	93	09
5	38	13	98	28

The ANOVA would assign different F-values to each of these five combinations. These F-values, therefore, represent the scores for each of the different customers. Generally, the higher the F-value, the more distinct a statement is among groups, and therefore it is considered to be a "higher" score. The customers can then be ranked according to their respective F-values.

In addition to the combinatorial methods described in steps **S110** and **S112**, other statistical combinations of the ranks from each segmentation strategy can be used, in step **S114**, to generate a composite score (and rank) for each customer.

In a preferred embodiment, the results from the above two methods for consolidating different segmentation strategy results, or other exemplary methods, can themselves be combined. To combine, for example, the results from the above two methods, the ranks for each customer are calculated using the two different methods and then averaged to generate, in step **S116**, an overall score for each customer. Again, the average could be weighted in consideration of other factors, for example, expert judgement. The customers can then be ranked, in step **S118**, according to their overall score.

Alternatively, the overall rankings generated in step **S118** can be determined from the scores of only one composite method (e.g., steps **S110**, **S112**, and **S114**) or from a different combination of the composite methods than described above as the preferred embodiment.

From the ranked list of customers, a specific portion of the top ranks are selected, in step **S120**, to receive marketing materials. Selection of the specific portion of the top ranks could depend on factors such as a desired response rate, a desired dollar value per transaction, average revenues received per marketing contact, or necessary quantity considerations.

As mentioned earlier, conventional marketing wisdom conforms to the belief that to optimize performance only a single segmentation strategy should be created to explain a single desired outcome (for example, a desired outcome could be "respond to a specific marketing program"). The precepts of the present invention, however, are diametrically opposed to the conventional wisdom in that they provide for the simultaneous use of multiple segmenting strategies aimed at multiple desired outcomes. This novel approach to segmentation provides results that consistently surpass the methods that conform to the conventional wisdom. The customers thus selected to receive marketing materials have a far greater likelihood of maximizing the profitability of the particular marketing program than any potential subset of customers that would have been selected using only a single segmentation strategy.

FIG. 2 is a block diagram that illustrates a computer system **100** upon which an embodiment of the invention may be implemented. Computer system **100** includes a bus **102** or other communication mechanism for communicating information, and a processor **104** coupled with bus **102** for processing information. Computer system **100** also includes a main memory **106**, such as a random access memory (RAM) or other dynamic storage device, coupled to bus **102** for storing information and instructions to be executed by processor **104**. Main memory **106** also may be used for storing temporary variables or other intermediate information during execution of instructions to be executed by processor **104**. Computer system **100** further includes a read only memory (ROM) **108** or other static storage device coupled to bus **102** for storing static information and

2010 IEEE. All rights reserved.

instructions for processor 104. A storage device 110, such as a magnetic disk or optical disk, is provided and coupled to bus 102 for storing information and instructions.

Computer system 100 may be coupled via bus 102 to a display 112, such as a cathode ray tube (CRT), for displaying information to a computer user. An input device 114, including alphanumeric and other keys, is coupled to bus 102 for communicating information and command selections to processor 104. Another type of user input device is cursor control 116, such as a mouse, a trackball, or cursor direction keys for communicating direction information and command selections to processor 104 and for controlling cursor movement on display 112. This input device typically has two degrees of freedom in two axes, a first axis (e.g., x) and a second axis (e.g., y), that allows the device to specify positions in a plane.

Computer system 100 operates in response to processor 104 executing one or more sequences of one or more instructions contained in main memory 106. Such instructions may be read into main memory 106 from another computer-readable medium, such as storage device 110. Execution of the sequences of instructions contained in main memory 106 causes processor 104 to perform the process steps described herein. In alternative embodiments, hard-wired circuitry may be used in place of or in combination with software instructions to implement the invention. Thus, embodiments of the invention are not limited to any specific combination of hardware circuitry and software.

The term "computer-readable medium" as used herein refers to any medium that participates in providing instructions to processor 104 for execution. Such a medium may take many forms, including but not limited to, non-volatile media, volatile media, and transmission media. Non-volatile media includes, for example, optical or magnetic disks, such as storage device 110. Volatile media includes dynamic memory, such as main memory 106. Transmission media includes coaxial cables, copper wire and fiber optics, including the wires that comprise bus 102. Transmission media can also take the form of acoustic or light waves, such as those generated during radio-wave and infra-red data communications.

Common forms of computer-readable media include, for example, a floppy disk, a flexible disk, hard disk, magnetic tape, or any other magnetic medium, a CD-ROM, any other optical medium, punchcards, papertape, any other physical medium with patterns of holes, a RAM, a PROM, and EPROM, a FLASH-EPROM, any other memory chip or cartridge, a carrier wave as described hereinafter, or any other medium from which a computer can read.

Various forms of computer readable media may be involved in carrying one or more sequences of one or more instructions to processor **104** for execution. For example, the instructions may initially be carried on a magnetic disk of a remote computer. The remote computer can load the instructions into its dynamic memory and send the instructions over a telephone line using a modem. A modem local to computer system **100** can receive the data on the telephone line and use an infra-red transmitter to convert the data to an infra-red signal. An infra-red detector can receive the data carried in the infra-red signal and appropriate circuitry can place the data on bus **102**. Bus **102** carries the data to main memory **106**, from which processor **104** retrieves and executes the instructions. The instructions received by main memory **106** may optionally be stored on storage device **110** either before or after execution by processor **104**.

Computer system **100** also includes a communication interface **118** coupled to bus **102**. Communication interface **118** provides a two-way data communication coupling to a network link **120** that is connected to a local network **122**. For example, communication interface **118** may be an integrated services digital network (ISDN) card or a modem to provide a data communication connection to a corresponding type of telephone line. As another example, communication interface **118** may be a local area network (LAN) card to provide a data communication connection to a compatible LAN. Wireless links may also be implemented. In any such implementation, communication interface **118** sends and receives electrical, electromagnetic or optical signals that carry digital data streams representing various types of information.

Network link **120** typically provides data communication through one or more networks to other data devices. For example, network link **120** may provide a connection through local network **122** to a host computer **124** or to data equipment operated by an Internet Service Provider

(ISP) 126. ISP 126 in turn provides data communication services through the world wide packet data communication network now commonly referred to as the "Internet" 128. Local network 122 and Internet 128 both use electrical, electromagnetic or optical signals that carry digital data streams. The signals through the various networks and the signals on network link 120 and through communication interface 118, which carry the digital data to and from computer system 100, are exemplary forms of carrier waves transporting the information.

Computer system 100 can send messages and receive data, including program code, through the network(s), network link 120 and communication interface 118. In the Internet example, a server 130 might transmit a requested code for an application program through Internet 128, ISP 126, local network 122 and communication interface 118. The received code may be executed by processor 104 as it is received, and/or stored in storage device 110, or other non-volatile storage for later execution. In this manner, computer system 100 may obtain application code in the form of a carrier wave.

While this invention has been described in connection with what is presently considered to be the most practical and preferred embodiment, it is to be understood that the invention is not limited to the disclosed embodiment, but on the contrary, is intended to cover various modifications and equivalent arrangements included within the spirit and scope of the appended claims. The invention is capable of other and different embodiments and its several details are capable of modifications in various obvious respects, all without departing from the invention. Accordingly, the drawings and description are to be regarded as illustrative in nature, and not as restrictive.